Current Biology 16, 1726–1729, September 5, 2006 ©2006 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2006.07.051

Report

Binocular Coordination of the Eyes during Reading

Simon P. Liversedge,[1,4,*] Keith Rayner,[2] Sarah J. White,[1,5] John M. Findlay,[1] and Eugene McSorley[3]
[1] Department of Psychology
University of Durham
Durham DH1 3LE
United Kingdom
[2] Department of Psychology
University of Massachusetts
Amherst, Massachusetts 01003
[3] School of Psychology
University of Reading
Reading, Berkshire RG6 6AL
United Kingdom

Summary

Saccadic eye movements and fixations are the behavioral means by which we visually sample text during reading. Human oculomotor control is governed by a complex neurophysiological system involving the brain stem, superior colliculus, and several cortical areas [1, 2]. A very widely held belief among researchers investigating primate vision is that the oculomotor system serves to orient the visual axes of both eyes to fixate the same target point in space. It is argued that such precise positioning of the eyes is necessary to place images on corresponding retinal locations, such that on each fixation a single, nondiplopic, visual representation is perceived [3]. Vision works actively through a continual sampling process involving saccades and fixations [4]. Here we report that during normal reading, the eyes do not always fixate the same letter within a word. We also demonstrate that saccadic targeting is yoked and based on a unified cyclopean percept of a whole word since it is unaffected if different word parts are delivered exclusively to each eye via a dichoptic presentation technique. These two findings together suggest that the visual signal from each eye is fused at a very early stage in the visual pathway, even when the fixation disparity is greater than one character (0.29 deg), and that saccade metrics for each eye are computed on the basis of that fused signal.

Results and Discussion

Primates have frontally placed eyes that facilitate precise binocular coordination and orienting. A very prevalent assumption in the field of eye-movement research is that both eyes fixate the same point in space and that this is achieved through very precise control of both the eyes. According to this view, such precise coordination ensures that aspects of the visual scene fall on corresponding retinal locations, thereby permitting the perception of a single, unified, visual percept.

A good example of a series of human psychological processes that rely upon the sequential uptake of visual information provided via saccadic sampling is the sentence-comprehension system involved in reading. When we read, we make a series of fixations (typically between 150 and 350 ms in duration), where we visually sample the text on the page, each followed by a conjugate saccadic eye movement where both eyes rotate in the same direction in order that novel text be directly inspected and processed [5]. In addition, during a fixation, small, comparatively slow, vergence movements of the eyes have been reported [6]. Clearly, under normal viewing conditions, diplopia is not experienced during reading. These observations are consistent with the view that, through saccadic and vergence eye movements, the eyes are always brought to fixate the same point (letter) within a word. Consequently, the same, corresponding, retinal input occurs for each eye, thereby permitting the construction of a unified perceptual representation of the fixated word from binocular retinal input.

We conducted binocular eye-tracking experiments to investigate this widely held assumption that people align both visual axes on the same letter while reading text. We hypothesized that fixation points of the left and right eye should be disparate by less than one letter on all fixations (spatial resolution well within the limits of our eye-tracking system). Surprisingly, the results showed that individuals often show substantial fixation disparity and frequently do not fixate the same letter within a word with each eye [7–9]. All of the analyses presented here focus on the results for a single target word within a sentence (see details below). Overall, first fixations on the critical word were 1.1 (SD = 0.8) characters apart at the beginning of the fixation and 1.0 (SD = 0.7) characters apart at the end. On 42% of fixations, the eyes were disparate by more than a character (averaging 1.8 [SD = 0.6] characters at fixation onset, with that disparity being only slightly reduced through vergence movements to a magnitude of 1.7 [SD = 0.5] characters by the end of fixation). When the points of fixation were disparate, the lines of gaze were generally diverged (uncrossed) relative to the text (93% of fixations), as has also been reported in other visual situations [10], but occasionally converged (crossed) (7% of fixations). Our data clearly show that on a substantial proportion of fixations during reading, the eyes do not provide retinal inputs that correspond perfectly about the point of fixation.

Having demonstrated that fixation disparity occurs frequently during reading, this raises a further important theoretical question about how saccade metrics are programmed. We were able to address this question by taking accurate binocular eye recordings, thus

*Correspondence: s.p.liversedge@soton.ac.uk
[4] Present address: Department of Psychology, University of Southampton, Southampton SO17 1BJ, United Kingdom.
[5] Present address: School of Psychology, University of Leicester, Leicester LE1 9 HN, United Kingdom.

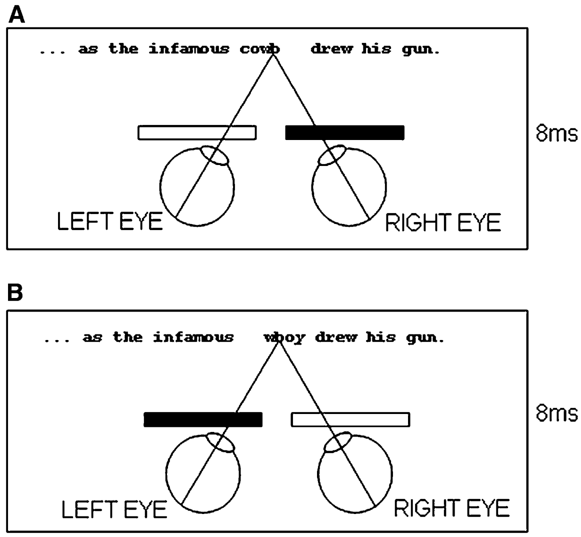

Figure 1. Schematic Representation of the Shutter Goggle Arrangement under Congruent Dichoptic Conditions to Each Eye

This arrangement shows the two alternating presentations of the target word *cowboy* in part of its sentential frame under congruent dichoptic conditions to each eye.

(A) The initial portion of the word was available to the left eye, but blocked to the right eye.

(B) The latter portion of the word was available to the right eye, but blocked to the left.

The two-letter overlap (*wb*) anchored the word halves in the vertical plane.

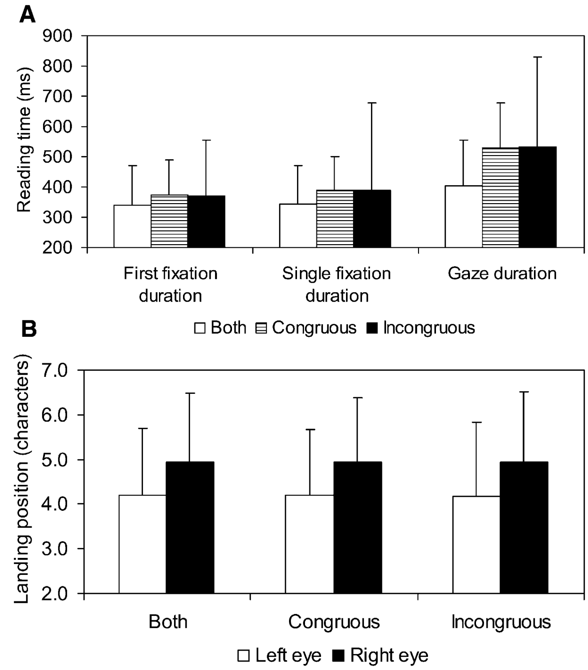

Figure 2. Mean Reading Times and Saccade Landing Positions on the Target Word for Each Condition

(A) Mean reading times (+1 standard deviation) for the target word under congruent and incongruent dichoptic viewing conditions, and the control condition where the whole word was presented to both eyes. First fixation duration is the duration of the first fixation on the word, regardless of whether the word was refixated. Single fixation duration is the duration of the first fixation on the word contingent on the participant not refixating the word. Gaze duration is the sum of all the fixations on the word until the participant made a saccade to another word.

(B) Mean saccadic landing positions (+1 standard deviation) on the target word for the left eye and the right eye under congruent and incongruent dichoptic viewing conditions, and the control condition where the whole word was presented to both eyes.

supplementing the numerous studies [5, 11] that have investigated eye-movement control during reading by recording the movements of just one eye. These studies have shown that upcoming words in the sentence that fall outside the fovea form targets for subsequent eye movements, and more specifically, saccades are targeted toward a position just to the left of center of the word.

Views concerning binocular coordination have continued to be polarized [12] after the classic debate in the early 19th century [13] between Hering, who advocated that binocular eye movements are yoked and driven by a single neural signal, and Helmholtz, who argued for independent programming of the eye movements of each eye although with strong crosscoupling. If a visually unified signal does drive binocular eye movements, then unification might be attained either through a process of suppression whereby one of the two inputs is blocked or through a process of fusion whereby inputs from the two eyes are combined, possibly assisted by some top-down visual processing. We used dichoptic presentation to test between these alternatives. To investigate whether saccades for each eye are programmed independently or on the basis of a unified perceptual representation, our experiment contained an additional manipulation. Within each sentence, we included a target word that was a compound noun (e.g., cowboy) that was 6, 8, or 10 letters long and was presented dichoptically (see Experimental Procedures and Figure 1), with constituent parts being presented separately to each eye. There were three presentation conditions: congruent (*cowb* to the left eye and *wboy* to the right eye); incongruent (*wboy* to the left eye and *cowb* to the right); and binocular control condition (*cowboy* to both eyes). If the input from one eye is suppressed, the saccade size should be dependent entirely on the visual input to the other eye and the landing positions would be entirely influenced by the input to the nonsuppressing eye. By contrast, if saccade metrics are computed on the basis of a fused representation of the word derived from different retinal signals, then mean fixation positions on the target word should be uninfluenced by the different dichoptic presentation conditions. Finally, if saccade metrics were computed for each eye independently on the basis of each eye's particular pattern of retinal stimulation, different saccade sizes would be expected in the two eyes.

The results showed that dichoptic presentation of word parts did significantly increase the duration of fixations on the target word relative to the control condition (see Figure 2A; all $Fs > 5$, $ps < .01$; all $ts > 2.6$, $ps < .05$). Also, clear differences between the landing position of the left and right eyes occurred, $F_1\ (1,14) = 52.21$, $p < 0.001$, $F_2\ (1,71) = 330.8$, $p < 0.001$, reflecting the basic uncrossed fixation disparity that occurs during reading that was reported earlier. However, critically, these landing positions on the target word were not affected by dichoptic presentation of the word (see Figure 2B; Fs < 1). The proportion of aligned, crossed, and uncrossed fixations made on the critical word was almost precisely the same under each of the different presentation conditions (control: aligned = .58, uncrossed = .4, crossed = .02; congruent: aligned = .58, uncrossed = .39, crossed = .03; incongruent: aligned = .60, uncrossed = .37, crossed = .03; Fs < 1).

These data indicate that despite the occurrence of binocular disparity during reading, saccade metrics for a nonfoveal target word are computed on the basis of a unified perceptual representation derived from the disparate retinal signals. Furthermore, these data demonstrate that the process by which the visual system attains a unified visual percept from disparate retinal signals is one of fusion rather than suppression. This suggests that during reading, the visual system combines disparate retinal signals from each eye at a very early stage of visual processing to produce a fused and unified visual percept of the word. Visual fusion is a well-established phenomenon but only when the retinal stimuli in each eye match to within a small region of correspondence. This region, known as Panum's fusional area, is usually held to be 5–10 min arc for stimuli with sharp edges [14]. Our results imply that in the active situation of reading, fusion occurs and tolerates a substantially greater mismatch (up to at least 30 min arc). It is on the basis of this percept that both the saccadic computation system and the human cognitive processing system operate to attain language comprehension. Together, these data illuminate psychological processes that are central to human binocular coordination that is necessary for successful visual processing to enable complex higher-order cognition during reading [11].

Experimental Procedures

Participants, Stimuli, and Apparatus

15 students at the University of Durham were paid to participate in the experiment. All participants had normal vision and were naive regarding the purpose of the experiment. The experimental sentences were all less than 73 characters long and included a target compound noun (e.g., cowboy) that was 6, 8, or 10 letters long. Different target-word lengths were included to ensure that effects were not length specific. Each of the two morphological constituents that comprised the compound noun were of equal length and each was itself a noun. Pseudocompound nouns (e.g., carpet) were not employed. Three lists of 77 items were constructed and comprised of 5 filler sentences and 72 experimental sentences, all with a variety of syntactic constructions. Conditions were rotated according to a Latin Square design. A series of repeated measures analyses of variance and t tests were undertaken across participant (F_1, t_1) and item (F_2, t_2) means. Sentences were displayed one at a time as white letters (in lower case except for where capital letters were appropriate) on a black background. 24 of the sentences were followed by a comprehension question to ensure that participants concentrated on understanding the sentences. Sentences were displayed on a Philips 21B582BH 24 in monitor at a viewing distance of 85 cm. Each character subtended 0.29° of visual angle. Movements of both eyes were monitored by left and right Fourward Technologies Dual Purkinje Image eye trackers. The resolution of the eye trackers is less than 10 min of arc and the sampling rate was every ms. The monitor and the eye trackers were interfaced with a Philips Pentium III PC that controlled the experiment. Dichoptic presentation of the target word was achieved through a pair of Cambridge Research Systems shutter goggles that block visual input to each eye alternately every 8 ms (corresponding to a 120 Hz refresh rate).

Procedure

Participants were told to read the sentences normally for comprehension. A bite bar and head restraint were used to minimize head movements. Prior to participants reading any sentences, the eye trackers were calibrated with a 9 point display. The calibration procedure is based on the values of the eye tracker output at each point and a set of linear interpolation routines that assign a horizontal and vertical eye direction measure to any eye tracker output values. During the left eye tracker calibration procedure, the right eye was occluded, and during the right eye calibration procedure, the left eye was occluded. Before the presentation of each sentence, the accuracy of each eye tracker was carefully checked and recalibrated whenever necessary. After reading each sentence, participants pressed a button to continue and used a button box to respond Yes/No to comprehension questions. The experiment lasted approximately 45 min.

Analyses

Eye movement records were analyzed with customized computer programs. Fixations were manually identified in order to avoid contamination by dynamic overshoots. 5% of trials were excluded due to tracker loss and 8% of the remaining trials were excluded due to the critical word being skipped on first pass. For the first fixation on the target word, cases were also excluded in which the fixation was under 80 ms or over 1200 ms (3% of fixations), the disparity was greater than 2.5 standard deviations above the mean for each participant (2% of fixations), or either the left or right eye was not fixating on the target (2% of fixations).

Acknowledgments

This research was supported by a grant from the Leverhulme Foundation (which supported K.R.'s stay at the University of Durham as a Visiting Leverhulme Professor), by Grant 12/S19168 from the Biotechnology and Biological Sciences Research Council, and by Grant HD26765 from the US National Institutes of Health. We thank Bob Metcalfe for his expert programming assistance.

Received: March 24, 2006
Revised: July 7, 2006
Accepted: July 7, 2006
Published: September 5, 2006

References

1. Wurtz, R.H. (1996). Vision for the control of movement: the Friedenwald lecture. Invest. Ophthalmol. Vis. Sci. *37*, 2131–2145.
2. Pierrot-Deseilligny, C., Müri, R.M., Ploner, C.J., Gaymard, B., Demeret, S., and Rivaud-Pecheux, S. (2003). Decisional role of the dorsolateral prefrontal cortex in oculomotor behaviour. Brain *126*, 1460–1473.
3. Howard, I.P., and Rogers, B.J. (1996). Binocular Vision and Stereopsis (New York: Oxford Univ. Press).
4. Findlay, J.M., and Gilchrist, I.D. (2003). Active Vision: The Psychology of Looking and Seeing (New York: Oxford Univ. Press).
5. Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychol. Bull. *124*, 372–422.
6. Hendriks, A.W. (1996). Vergence eye movements during fixations in reading. Acta Psychol. (Amst.) *92*, 131–151.
7. Juhasz, B.J., Liversedge, S.P., White, S.J., and Rayner, K. (2006). Binocular coordination of the eyes during reading: Word frequency and case alternation affect fixation duration but not binocular disparity. Q. J. Exp. Psychol. *59*, 1614–1625.
8. Liversedge, S.P., White, S.J., Findlay, J.M., and Rayner, K. (2006). Binocular coordination of eye movements during reading. Vision Res. *46*, 2363–2374.
9. Blythe, H.I., Liversedge, S.P., Joseph, H.S.S.L., White, S.J., Findlay, J.M., and Rayner, K. (2006). The binocular coordination of eye movements during reading in adults and children. Vision Res., in press. Published online August 1, 2006. 10.1016/j.visres.2006.06.006.
10. Erkelens, C.J., Steinman, R.M., and Collewijn, H. (1989). Ocular vergence under natural conditions. II Gaze shifts between real

targets differing in distance and direction. Proc. R. Soc. Lond. B Biol. Sci. *236B*, 441–465.
11. Liversedge, S.P., and Findlay, J.M. (2000). Saccadic eye movements and cognition. Trends Cogn. Sci. *4*, 6–14.
12. Wade, N.J., and Tatler, B.W. (2005). The Moving Tablet of the Eye (Oxford: Oxford Univ. Press).
13. King, W.M., and Zhou, W. (2000). New ideas about binocular coordination of eye movements: is there a chameleon in the primate tree? Anat. Rec. *261*, 153–161.
14. Howard, I.P., and Rogers, B.J. (1995). Binocular Vision and Stereopsis (New York: Oxford Univ. Press).